CS6705 FOUNDATIONS OF ARTIFICIAL INTELLIGENCE

Introduction

CS6705

- http://www.cs.unb.ca/~ulieru/index.htm
- Please download Course Outline
- Textbook: S. Russell and P. Norvig *Artificial Intelligence: A Modern Approach* Prentice Hall, 2003, Second Edition
- Lecturer: Mihaela Ulieru

Outline

- Course overview
- What is AI?
- A brief history
- The state of the art

Scientific / philosophical goals of AI

- Understand cognitive capabilities of biological organisms
  - Brains, neurons, neurotransmitters, …
- Resolve mind/body dualism
  - Consciousness, determinism, mental vs physical worlds

Engineering goals of AI

- Develop tools that can mimic capabilities of “intelligent” entities (e.g. humans)
  - Sensory responsiveness
    - Speech, vision, …
  - Manipulation (control) of environment
    - Planning, actions, feedback
  - Complex data processing
    - Retrieval, filtering, creativity

Approaches to AI

- Symbolic AI:
  - *Intelligence*: manipulation of symbols
  - *AI*: symbolic representation of knowledge, search through representations
- Statistical AI:
  - *Intelligence*: learn from statistical regularities in data
  - *AI*: development of “information filters”
- Embodied AI:
  - *Intelligence*: dealing with challenges in environment
  - *AI*: construction of agents that combine “rational” & “emotive” capabilities

The symbolic approach ("classical AI")

"Yes, yes, I *know* that, Sidney… *every*body knows *that!*… But look: Four wrongs *squared,* minus two wrongs to the fourth power, divided by this formula, *do* make a right."

Knowledge representation

- Goal: develop data structures that make it easy to represent diverse, “real-world” knowledge
- Typical approaches
  - Logic (propositional, predicate, HO)
  - Semantic networks (nodes, named links, hierarchies)
- Major challenge is automation of knowledge capture (“knowledge engineering”)

What is AI?

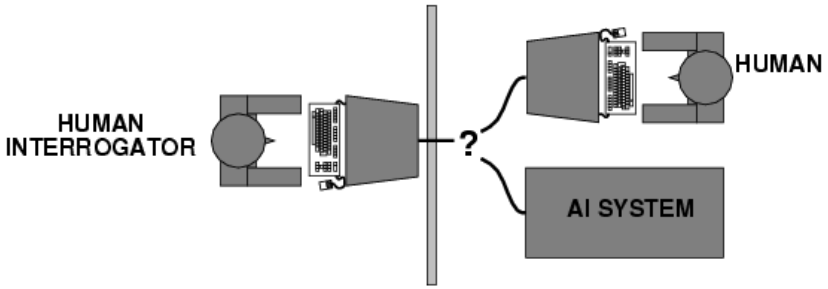
Views of AI fall into four categories:

Thinking humanly	Thinking rationally
Acting humanly	Acting rationally

The textbook advocates "acting rationally"

Acting humanly: Turing Test

- Turing (1950) "Computing machinery and intelligence":
- "Can machines think?" → "Can machines behave intelligently?"
- Operational test for intelligent behavior: the Imitation Game

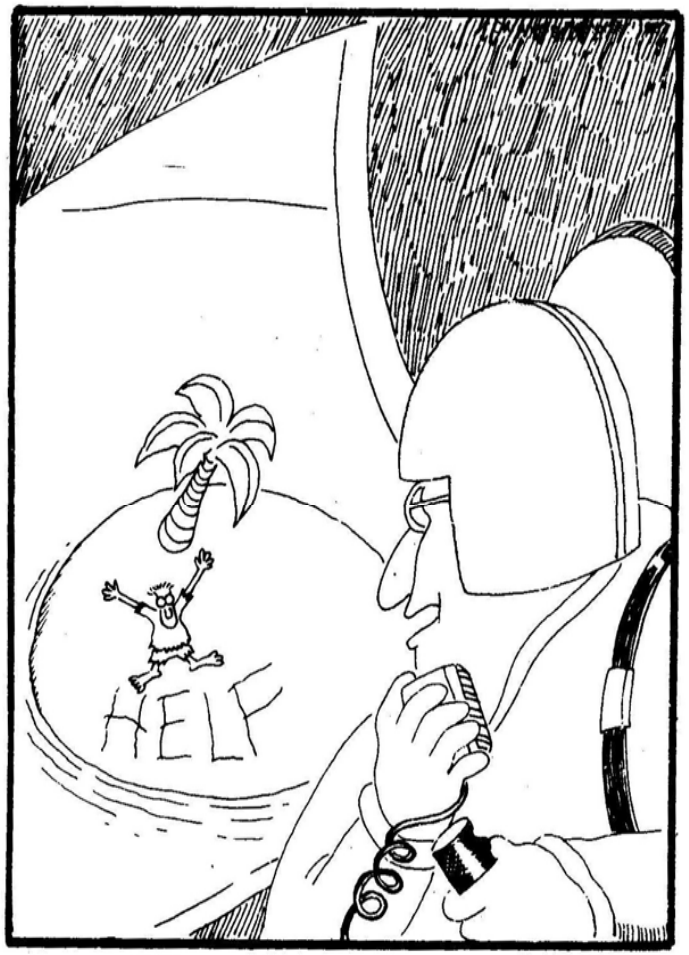


- Predicted that by 2000, a machine might have a 30% chance of fooling a lay person for 5 minutes
- Anticipated all major arguments against AI in following 50 years
- Suggested major components of AI: ***knowledge, reasoning, language understanding, learning***

Thinking humanly: cognitive modeling

- 1960s "cognitive revolution": information-processing psychology
- Requires scientific theories of internal activities of the brain
- -- How to validate? Requires
  1) Predicting and testing behavior of human subjects (top-down)
  or 2) Direct identification from neurological data (bottom-up)
- Both approaches (roughly, Cognitive Science and Cognitive Neuroscience)
- are now distinct from AI

Embodied agents



"Wait! Wait! . . . Cancel that, I guess it says 'helf.'"

Emotions and AI

- Growing awareness: role of
  - Environment
  - Resource constraints
- Emotions encapsulate relevant information
  - Contribute to intelligent behavior
  - Support “social” interaction

Computational models of emotions

- Ethology-inspired models
  - *Model agents, environment, survival constraints*
- Emotion-related learning
  - *Provide feedback (reinforcement) on behaviors to guide learning*
- Appraisal-based models
  - *Features of situation map to emotional states*
- Architecture-level models
  - *Combination of mechanisms to support successful behavior (e.g. Sloman's CogAff)*

Thinking rationally: "laws of thought"

- Aristotle: what are correct arguments/thought processes?
- Several Greek schools developed various forms of *logic*: *notation* and *rules of derivation* for thoughts; may or may not have proceeded to the idea of mechanization
- Direct line through mathematics and philosophy to modern AI
- Problems:
  1. Not all intelligent behavior is mediated by logical deliberation
  2. What is the purpose of thinking? What thoughts should I have?

Acting rationally: rational agent

- Rational behavior: ***doing the right thing***
- The right thing: that which is expected to ***maximize goal achievement***, given the available information
- Doesn't necessarily involve thinking – e.g., blinking reflex – but thinking should be in the service of rational action

Rational agents

- An agent is an entity that perceives and acts
- This course is about designing rational agents
- Abstractly, an agent is a function from percept histories to actions:

$$[f: \mathcal{P}^{\star} \rightarrow \mathcal{A}]$$

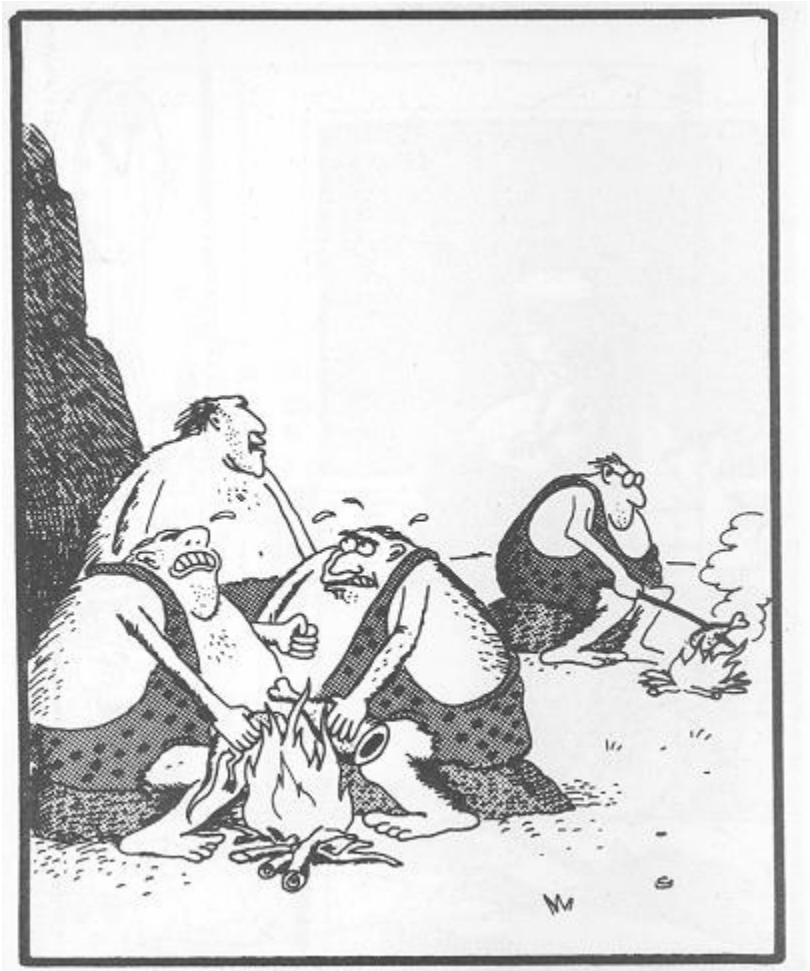
- For any given class of environments and tasks, we seek the agent (or class of agents) with the best performance
- Caveat: computational limitations make perfect rationality unachievable

  → design best program for given machine resources

AI prehistory

• Philosophy	Logic, methods of reasoning, mind as physical system foundations of learning, language, rationality
• Mathematics	Formal representation and proof algorithms, computation, (un)decidability, (in)tractability, probability
• Economics	utility, decision theory
• Neuroscience	physical substrate for mental activity
• Psychology	phenomena of perception and motor control, experimental techniques
• Computer engineering	building fast computers
• Control theory	design systems that maximize an objective function over time
• Linguistics	knowledge representation, grammar

Looking back, looking ahead



"Hey! Look what Zog do!"

A history of overselling

- *“Within ten years a digital computer will be the world's chess champion*”

A history of overselling

- *“Within ten years a digital computer will be the world's chess champion*”
  - Newell, 1957

A history of overselling

- *“Within ten years a digital computer will be the world's chess champion*”
  - Newell, 1957
- *“There are now in the world machines that think, that learn and create”*

A history of overselling

- *“Within ten years a digital computer will be the world's chess champion*”
  - Newell, 1957
- *“There are now in the world machines that think, that learn and create”*
  - Simon & Newell, 1958

Achievements

- Grandmaster-level chess (and valued competitor in computer games)
- Standard part of logistics planning
- Telephone agent; security inspector; credit verifier
- Google!

The pessimistic view

- Computers will never match human intelligence because:
  - They are not “grounded” (*Searle*, *Dreyfus*)
  - They lack the appropriate quantum-mechanical machinery (*Penrose*)
  - Dualism was right after all (*Chalmers*)

The optimistic view (1)

- Tasks that have been considered *intelligent*:
  - Solving algebraic problems
  - Playing chess
  - Understanding speech
- AI keeps moving the frontier!
- Incremental approach will take us past *X*, for any *X*

What next?

- Continued uptake of “AI” methods & tools in computing
  - Search, relational methods, neural networks → Bayesian networks, agent-based processing
- Growth of AI as more tasks must be automated
  - Internet bots, ubiquitous networks, security networks, 24 x 7 customer service, …
- A return to fundamental research

Abridged history of AI

- 1943 McCulloch & Pitts: Boolean circuit model of brain
- 1950 Turing's "Computing Machinery and Intelligence"
- 1956 Dartmouth meeting: "Artificial Intelligence" adopted
- 1952—69 Look, Ma, no hands!
- 1950s Early AI programs, including Samuel's checkers program, Newell & Simon's Logic Theorist, Gelernter's Geometry Engine
- 1965 Robinson's complete algorithm for logical reasoning
- 1966—73 AI discovers computational complexity
  Neural network research almost disappears
- 1969—79 Early development of knowledge-based systems
- 1980-- AI becomes an industry
- 1986-- Neural networks return to popularity
- 1987-- AI becomes a science
- 1995-- The emergence of intelligent agents

ASIGNMENTS

- 1. Read Chapter 1 in your book
- 2. Pick your favorite question at the end of Chapter 1 and write a min half page, max one page essay with your answer. Reflect on what is of utmost interest to you (e.g. relative to your quest in your graduate studies, or your personal / hobby interest) from the field of AI – and how can you use this class to dive into this particular field.